



**FACULTAD DE CIENCIAS
ECONÓMICAS Y EMPRESARIALES**

**GRADO EN ADMINISTRACIÓN Y
DIRECCIÓN DE EMPRESAS
TRABAJO DE FIN DE GRADO**

TÍTULO: *La transformación Box-Cox*

AUTOR: *Ana Verónica de la Guía González*

TUTOR/ES: *David Casado de Lucas*

Contenido

1. Introducción	2
2. Las transformaciones	3
2.1 Introducción.....	3
2.2 La transformación Box-Cox	8
3. Aplicaciones de la transformación Box-Cox	13
3.1 Una variable.....	13
3.2 Dos variables.....	16
4. Ejemplos y ejercicios de simulación.....	18
5. Conclusiones	28
6. Anexos.....	31
7. Referencias.....	36

1. Introducción

La estadística a menudo se enfrenta con datos que no presentan las características deseadas o requeridas para la situación en la que se está trabajando. Por ejemplo, que no sean simétricos, no estén centrados en el origen o porque no sigan una distribución normal. Una de las posibles soluciones para mejorar el análisis estadístico es el uso de transformaciones de datos (Sakia, 1992). Si pensamos en las variables generadoras de esos datos, se hablaría de transformación de variables.

Las distintas situaciones, estadísticamente hablando, se caracterizan por un número de variables, el problema que se quiere resolver y, por último, el método para resolverlo. En referencia al número de variables, en unas ocasiones se desea describir una sola variable. En otras, se pretende estudiar la relación entre dos o más variables, etc.

Las transformaciones se pueden aplicar a una sola variable X , a dos variables (X, Y) , o a más variables. En el primer caso, las razones para llevar a cabo una transformación son principalmente las siguientes: modificar la media o la varianza, reducir la asimetría o la curtosis, aunque especialmente se hace para tratar de conseguir que los datos sigan una distribución normal. Transformar una variable consiste, en notación matemática en aplicar $f(X)=Y$, a la variable X .

El estudio de dos variables (X, Y) viene motivado, entre otras razones, por intentar ajustar la relación entre ambas. En algunos problemas económicos aplicados, se parte de la base de una variable dada X , independiente o exógena cuyos valores son determinados por factores o actividades realizadas fuera del sistema económico examinado, y una variable dependiente Y cuyo valor viene determinado por X . En este caso, el problema de estudio consistiría en analizar $g(X, Y)=0$, que casi siempre se estudia en la forma explícita $Y=f(X)$, donde es importante resaltar que tanto X como Y son conocidas. Un caso típico sería la regresión lineal, en la cual es frecuente buscar una transformación en la que la relación entre las variables transformadas sea lineal, que la variable dependiente y el error se distribuyan siguiendo una distribución normal, sin heterocedasticidad ni autocorrelación.

En este trabajo, se ha abordado la familia de transformaciones no lineales propuestas por los autores George Box y David Cox en 1964, el objetivo y la importancia de dichas transformaciones, así como una aproximación teórica y, por último, se ha complementado con ejemplos.

En primer lugar se ha realizado una pequeña introducción de los distintos tipos de transformaciones, incluyendo las lineales y las no lineales. A continuación, se ha profundizado en las transformaciones Box-Cox, las fórmulas matemáticas y los motivos que justifican llevar a cabo una transformación de este tipo. Además, se han separado las transformaciones de una sola variable, de aquellas de varias variables. Por último, se han incluido ejemplos para mostrar al lector la aplicación dichas transformaciones.

2. Las transformaciones

2.1 Introducción

Dependiendo del objetivo que se pretenda conseguir, de las características de los datos, y de la técnica estadística, será conveniente aplicar un tipo de transformación concreta u otro. En este trabajo se introducirán las transformaciones lineales y a continuación se hará un análisis más exhaustivo de las no lineales, centrándose sobre todo en la familia de transformaciones Box-Cox.

Las transformaciones lineales,

$$ax + b = y$$

cambian valores numéricos de la variable pero no modifican ni la asimetría ni la curtosis de la distribución (Osborne, 2002; Fornieles, 2013; McCune y Grace, 2002). El lector interesado puede encontrar la demostración en el Anexo 1 incluido al final del trabajo. Ejemplos particulares de transformaciones lineales son la traslación ($a=1$) y la homotecia ($b=0$).

Un ejemplo clásico de transformación lineal es la tipificación de datos. Esta resulta muy útil para eliminar la dependencia con respecto a las diferentes unidades de medidas

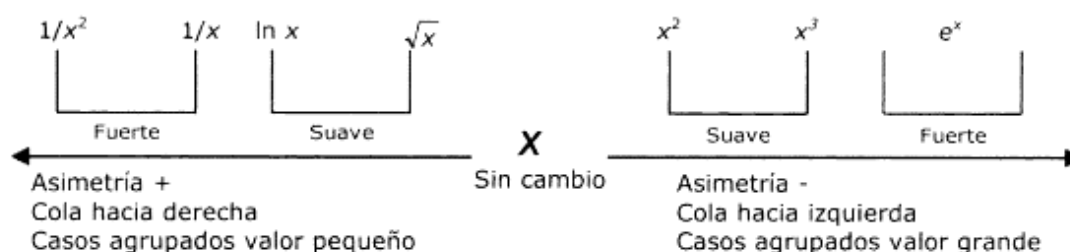
empleadas (Lane, 2010). Por ejemplo, el cambio de unidades de medida como cuando se pasa de grados Fahrenheit a grados Celsius.

$$F = \left(\frac{9}{5} * C\right) + 32$$

Las transformaciones no lineales, por el contrario, permiten corregir la forma de la distribución, por ejemplo la asimetría y la curtosis (Fornieles, 2013). Dentro del grupo de las no lineales se encuentra la familia de transformaciones Box-Cox.

El origen de estas transformaciones se remonta a la escalera de Tukey, la cual incluye una serie de transformaciones por potencias (Sakia, 1992). En ella además se muestra el tipo de transformación recomendada según el tipo de distribución, es decir, atendiendo a su asimetría o la dirección en la que se encuentran los casos más extremos. Erickson y Nosanchuk resumieron la escalera de Tukey en el siguiente gráfico (Guisande González, 2006):

Figura: 1 Escalera de Tukey (Fuente: Guisande González, 2006).



Tal y como se aprecia en la Figura 1, si la distribución es asimétrica positiva, será conveniente utilizar transformaciones del tipo raíces cuadradas o logarítmicas. La corrección de la asimetría positiva será aún mayor con los logaritmos (Guisande González, 2006). Por el contrario, si se trata de una asimetría negativa, será conveniente elevar al cuadrado o al cubo.

Es importante destacar que estas transformaciones suelen buscar mejorar la normalidad, modificando las distancias relativas entre los puntos. Será entonces cuando se pueda interpretar adecuadamente los resultados (Osborne, 2002).

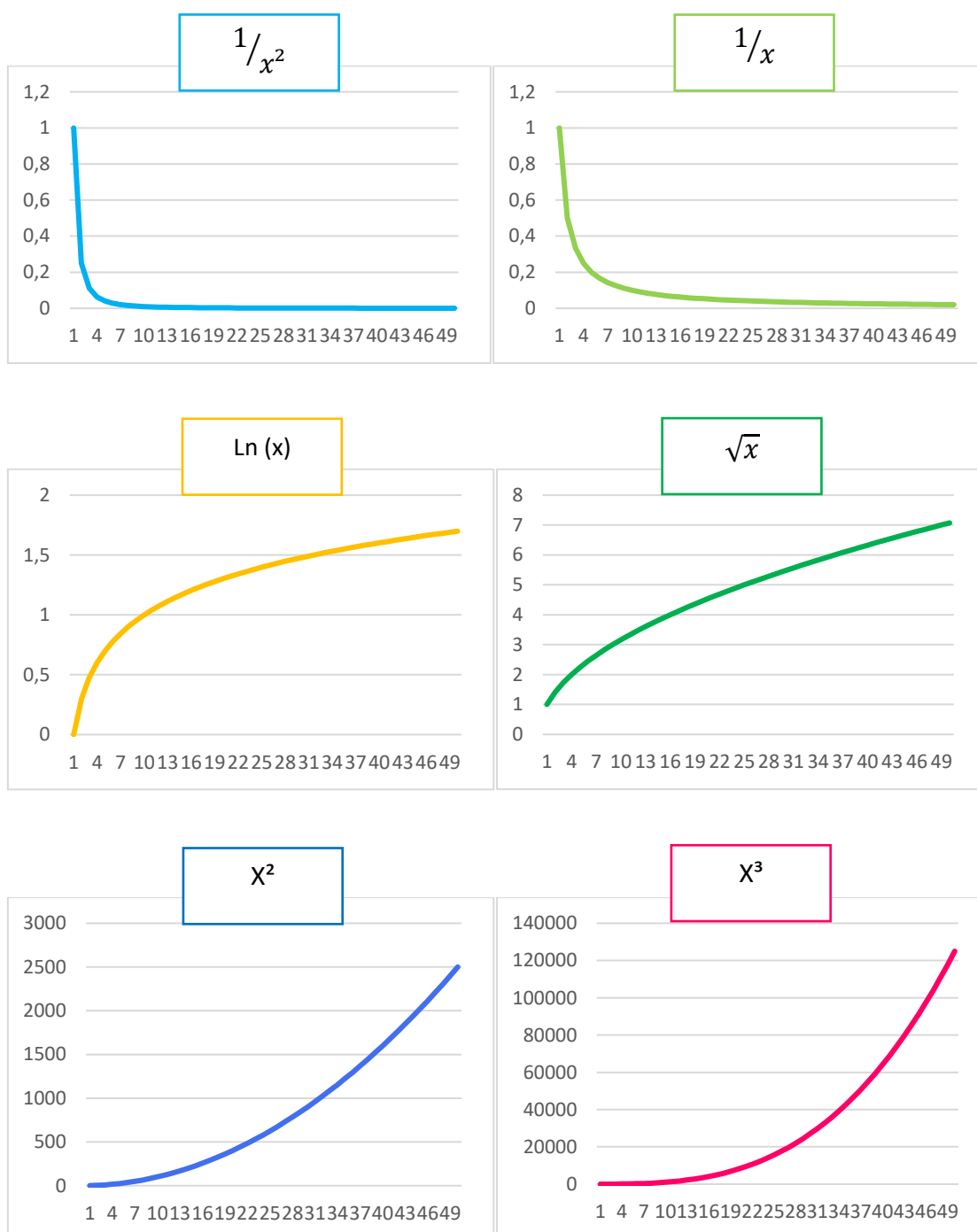
Aunque se estudiarán más adelante las propiedades matemáticas principales de la transformación de Box-Cox, que incluye a estas exponenciales, conviene mencionar que la corrección de los datos será mayor cuanto mayor es la exponenciación (Guisande González, 2006), y que si se aplica una transformación monótona creciente, todos los puntos continuarán en el mismo orden que antes de realizar la transformación (Osborne, 2002).

En el trabajo se considera que los datos son positivos puesto que el logaritmo solo está definido para estos valores. Los datos no positivos siempre pueden ser trasladados hasta que lo sean.

Tabla 1 El efecto de las transformaciones en variables (Fuente: Osborne, 2002).

Original Y	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00
SquareRoot(Y)	1.00	1.41	1.73	2.00	2.24	2.45	2.65	2.83	3.00	3.16
gap		0.41	0.32	0.27	0.24	0.21	0.20	0.18	0.17	0.16
% reduction	0.00	29.29	42.26	50.00	55.28	59.18	62.20	64.64	66.67	68.38
Log10 (Y)	0.00	0.30	0.48	0.60	0.70	0.78	0.85	0.90	0.95	1.00
gap		0.30	0.18	0.12	0.10	0.08	0.07	0.06	0.05	0.05
% reduction	100.00	84.95	84.10	84.95	86.02	87.03	87.93	88.71	89.40	90.00
Reflected Inverse(Y)	0.00	0.50	0.67	0.75	0.80	0.83	0.86	0.88	0.89	0.90
gap		0.50	0.17	0.08	0.05	0.03	0.02	0.02	0.01	0.01
% reduction	100.00	75.00	77.78	81.25	84.00	86.11	87.76	89.06	90.12	91.00
Original Y	11.00	12.00	13.00	14.00	15.00	16.00	17.00	18.00	19.00	20.00
SquareRoot(Y)	3.32	3.46	3.61	3.74	3.87	4.00	4.12	4.24	4.36	4.47
gap		0.15	0.14	0.14	0.13	0.13	0.12	0.12	0.12	0.11
% reduction	69.85	71.13	72.26	73.27	74.18	75.00	75.75	76.43	77.06	77.64
Log10 (Y)	1.04	1.08	1.11	1.15	1.18	1.20	1.23	1.26	1.28	1.30
gap		0.04	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.02
% reduction	90.53	91.01	91.43	91.81	92.16	92.47	92.76	93.03	93.27	93.49
Reflected Inverse(Y)	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95
gap		0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
% reduction	91.74	92.36	92.90	93.37	93.78	94.14	94.46	94.75	95.01	95.25
Original Y	100.00	101.00	102.00	103.00	104.00	105.00	106.00	107.00	108.00	109.00
SquareRoot(Y)	10.00	10.05	10.10	10.15	10.20	10.25	10.30	10.34	10.39	10.44
gap		0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
% reduction	90.00	90.05	90.10	90.15	90.19	90.24	90.29	90.33	90.38	90.42
Log10 (Y)	2.00	2.00	2.01	2.01	2.02	2.02	2.03	2.03	2.03	2.04
gap		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% reduction	98.00	98.02	98.03	98.05	98.06	98.08	98.09	98.10	98.12	98.13
Reflected Inverse(Y)	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
gap		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% reduction	99.01	99.02	99.03	99.04	99.05	99.06	99.07	99.07	99.08	99.09

Figura 2: Efecto de las transformaciones. (Fuente: elaboración propia).

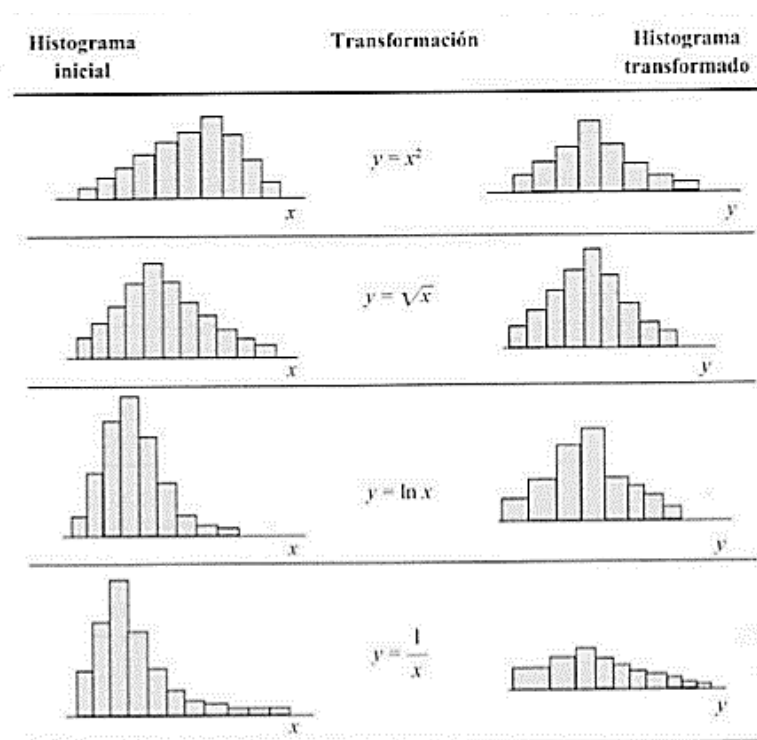


En la Figura 2, se representa en el eje de las X el valor previo a la transformación, mientras que en el Y el nuevo valor. El lector puede utilizar estos gráficos para entender el efecto de estas transformaciones. Tal y como se puede apreciar en la Tabla 1, los nuevos valores tienen distancias relativas menores en los cuatro primeros gráficos. Por ejemplo, al aplicar la raíz cuadrada, los valores pasan de ser {1, 2, 3, 4} a {1, 1,41, 1,73,

2}, y al aplicar una transformación logarítmica {1, 2, 3, 4} se convierten en {0, 0,30, 0,48, 0,60} tal y como se aprecia en la Tabla 1. Sin embargo, en las dos últimas gráficas sucede exactamente lo opuesto. En el quinto gráfico los valores {1, 2, 3, 4} pasan a ser {1, 4, 9, 16} aumentando así la distancia relativa entre los puntos.

Según el autor Osborne (2002) las transformaciones recomendadas con mayor frecuencia suelen ser la raíz cuadrada, el logaritmo o la inversa. Tal y como se podrá apreciar posteriormente en el trabajo, las tres forman parte de la familia de transformaciones Box-Cox. En el caso concreto de la raíz cuadrada el autor sugiere que, antes de realizar una transformación de este tipo, se trasladen los datos de tal manera que el valor mínimo sea 1 porque en el intervalo $(1, \infty)$ aumenta considerablemente la eficacia de la transformación. Además, en la transformación logarítmica las bases propuestas son 2, 10 y e . Para este ejemplo concreto anterior se ha utilizado 10. No obstante cuando no se especifique la base, se utiliza el logaritmo natural o neperiano.

Figura 3: Efecto de las transformaciones (Fuente: Peña, 1997).



En la Figura 3 se observan varios ejemplos de distribuciones no normales a las que se les ha aplicado una transformación. El resultado de esta aparece en la parte derecha, donde

se refleja que en todos los casos ha mejorado considerablemente la forma de la distribución.

Para terminar esta sección y aunque se presentarán ejemplos concretos más adelante, se mencionan algunos ejemplos de transformaciones:

- a) Se pretende estudiar el crecimiento del consumo de energía en diferentes países. Para ello conviene estudiar las diferencias entre dos instantes de tiempo diferentes $C_t - C_{(t-1)}$. Sin embargo resulta más útil para el estudio utilizar las diferencias relativas $\frac{C_t - C_{(t-1)}}{C_{(t-1)}}$. Una medida más adecuada, incluyendo una transformación no lineal sería el uso de logaritmos. El modelo translog se interpretaría como una aproximación de segundo orden a una forma funcional que se desconoce (Greene, 2002). Así, se podrán estudiar los efectos de segundo orden como son la elasticidad de sustitución.
- b) Además en el caso concreto de las series temporales, una transformación logarítmica es empleada normalmente para estabilizar la varianza.

2.2 La transformación Box-Cox

Como se ha mencionado, las transformaciones lineales no modifican las medidas de forma (asimetría y curtosis). Dado que las transformaciones no lineales si modifican estas medidas, son útiles para obtener nuevas variables con distribuciones más parecidas a la de la normal. La siguiente familia de funciones exponenciales, que es bien conocida, incluye transformaciones muy útiles:

$$x^\lambda \quad \lambda = \pm 1, \pm 2, \dots$$

Matemáticamente nada impide considerar exponentes no enteros, lo que enriquece enormemente la familia:

$$x^\lambda \quad \lambda \in \mathbb{R}$$

En regresión, por ejemplo, pueden transformarse una o más variables para resolver problemas de falta de normalidad o de heterocedasticidad. Box y Cox (1964) mostraron

que el parámetro λ puede estimarse a la vez que los coeficientes del modelo de regresión utilizando el Método de Máxima Verosimilitud (Montgomery et al., 2002).

La elección del parámetro λ , es decir, del miembro de la familia más apropiado para los objetivos que se persiguen, puede hacerse de diferentes formas. Hay definidos muchos métodos de ajuste de curvas que podrían utilizarse para ir probando distintos valores de λ y después elegir el más adecuado. Pero también hay métodos más teóricos para estimar λ , aunque con frecuencia llevan a problemas de optimización que deben resolverse con métodos computacionales y la ayuda del ordenador. En cualquier caso, la obtención de valores para λ cercanos a 0 supone un problema, puesto que $x^0 = 1$.

Aunque con aplicaciones que van más allá de la regresión, fue en este entorno en el que Box y Cox (1964) propusieron una familia de transformaciones que evitaba el problema mencionado y además podía resolver simultáneamente varios de los problemas que aparecen al ajustar modelos de regresión. Esta familia es:

$$x(\lambda) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log x & \text{if } \lambda = 0 \end{cases}$$

Donde x es la variable que se va a transformar y λ el parámetro de transformación. Dado que $\frac{x^\lambda - 1}{\lambda}$ tiende a $\log x$ cuando λ tiende a 0, se consigue la continuidad en λ y esto hace que la función $\log x$ quede 'incluida' en la familia de transformaciones. Por otra parte al ser crecientes todas las transformaciones de la familia, como ya veremos, se podrá conseguir la linealidad entre dos variables si la relación entre ellas es monótona (Peña, 1997). Esta familia de transformaciones tiene muchas aplicaciones, por ejemplo, en problemas con una sola variable o como método auxiliar para otras técnicas.

La transformación Box-Cox puede generalizarse de varias formas:

- Dado que el logaritmo solo está definido para número positivos, Box y Cox (1964) también introdujeron un segundo parámetro que permite trasladar los valores antes de aplicar la anterior expresión:

$$x(\lambda) = \begin{cases} \frac{(x + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & \text{if } \lambda_1 \neq 0 \\ \log(x + \lambda_2) & \text{if } \lambda_1 = 0 \end{cases}$$

En la práctica se elige λ_2 para que $x_i + \lambda_2 > 0$ para todo i , de manera que λ_1 es el único parámetro. Otros autores aconsejan trasladar de modo que el valor mínimo coincida con 1.

- Para distribuciones simétricas, John y Draper (1980) propusieron la siguiente familia:

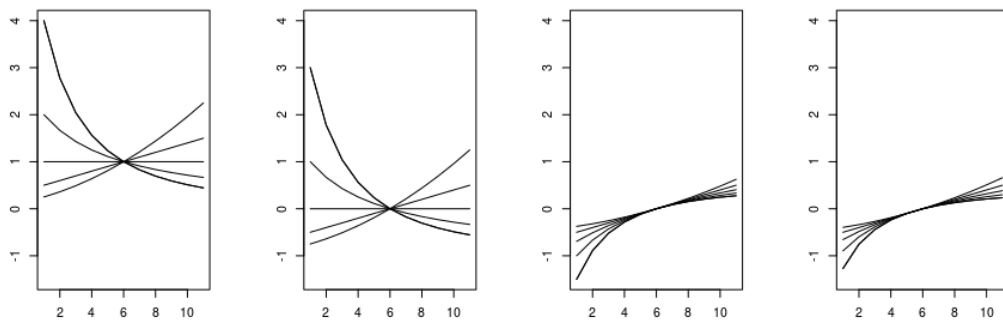
$$x(\lambda) = \begin{cases} \text{signo}(x) \frac{(|x| + 1)^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \text{signo}(x) \log(|x| + 1) & \lambda = 0 \end{cases}$$

- Para poder comparar los resultados de regresiones para distintos valores de λ , concretamente, las sumas de los cuadrados de los residuos, se utiliza:

$$x(\lambda) = \begin{cases} \frac{x^\lambda - 1}{\lambda \dot{x}^{\lambda-1}} & \lambda \neq 0 \\ \dot{x} \log(|x| + 1) & \lambda = 0 \end{cases}$$

donde $\dot{x} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$ es la media geométrica. Tiene la ventaja de que cambia más lentamente con λ (Montgomery et al., 2002).

Figura 4: Las cuatro familias de transformaciones y sus efectos. (Fuente: elaboración propia).



En la Figura 4 se muestra las transformaciones de las distintas familias descritas anteriormente. Para realizar estos gráficos se ha utilizado el programa estadístico R, cuyo código está en el Anexo 2.

Una de las aplicaciones más importante de esta familia de transformaciones consiste en obtener datos que sigan una distribución lo más parecida a la normal. Aunque no siempre se garantiza que después de haber hecho la transformación los nuevos datos

sigan una distribución normal, sí que se reducirán los problemas relacionados con la estimación, predicción e interferencia. Es preciso, por tanto, tras haber aplicado las transformaciones, utilizar algún concepto estadístico para evaluar si efectivamente se ha mejorado la distribución o se ha conseguido la normalidad.

Para estimar el parámetro λ , los autores Box y Cox (1964) propusieron fundamentalmente dos formas: el Método de Máxima Verosimilitud y el Método Bayesiano. En el primer caso, se pretende encontrar un valor de λ que minimice la suma de cuadrados de los residuos. Con frecuencia, aunque no siempre se espera que los valores de λ vayan de -2 a 2.

La estrategia de probar distintas transformaciones de los datos (o variables) y evaluar los resultados ha sido utilizada para encontrar variables transformadas que se relacionan linealmente, lo que equivale a encontrar una relación (no lineal) entre las variables originales. A esta búsqueda se la denomina ‘empírica’.

Las siguientes transformaciones son las más típicas:

Tabla 2: Transformaciones Box-Cox más típicas. (Fuente: Elaboración propia).

λ	X'
-2	$1/X^2$
-1	$1/X$
-0,5	$1/\sqrt{X}$
0	Log (X)
0,5	\sqrt{X}
1	X
2	X^2

La Tabla 2 recoge un resumen de las transformaciones más utilizadas. A la izquierda se muestran los posibles valores de λ mientras que a la derecha se incluye la expresión matemática de la transformación correspondiente. A lo largo del trabajo se hará referencia en varias ocasiones a esta tabla.

Una forma de comprender mejor la transformación de Box-Cox consiste en estudiar sus propiedades matemáticas básicas. Para ello, es además útil seguir el mismo camino que se ha seguido al introducir esta familia de transformaciones. Para valores positivos ($x > 0$) se cumplen las siguientes propiedades:

- x^λ , $\lambda = \pm 1, \pm 2, \dots$ es continua y creciente en x . Como λ es discreto, no tiene sentido hablar de continuidad en λ .
- x^λ , $\lambda \in \mathbb{R}$ es continua y creciente en x y además continua en λ . Para exponentes no enteros se puede escribir $x^\lambda = e^{\lambda \log(x)}$.
- $x^\lambda - 1$, $\lambda \in \mathbb{R}$ es una traslación vertical de la familia anterior, por lo que se preservan sus propiedades. Como x^λ pasa por el punto (1,1), esta familia trasladada pasa por el punto (1,0).
- $\frac{x^\lambda - 1}{\lambda}$, $\lambda \in \mathbb{R} \setminus \{0\}$ es una homotecia sobre la familia anterior, que puede escribirse también como $\frac{1}{\lambda}(x^\lambda - 1)$. Por tanto, no cambian las anteriores propiedades de continuidad en x . Como λ puede ser negativo, no es evidente que se preserve el crecimiento, lo que se demuestra mirando a su derivada:

$$\frac{1}{\lambda} \lambda x^{(\lambda-1)} = x^{(\lambda-1)} > 0 \quad (\text{si } x > 0)$$

Por último, como función de λ es el cociente de dos funciones continuas, por lo que es también continua si $\lambda \neq 0$. Como se ve, hay una discontinuidad cuando $\lambda = 0$. Para ver si es evitable y por tanto es posible definir ('rellenar') la función en ese punto, se resuelve el límite $0/0$ con la regla de L'Hôpital:

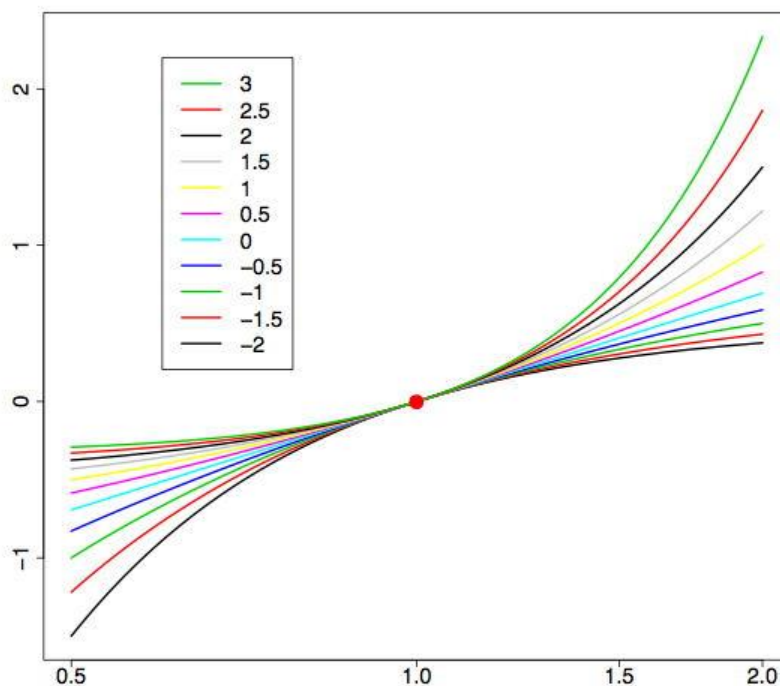
$$\lim_{\lambda \rightarrow 0} \frac{x^\lambda - 1}{\lambda} = \lim_{\lambda \rightarrow 0} \frac{x^\lambda \log(x)}{1} = \log(x)$$

La Figura 5 muestra el efecto de las transformaciones para distintos valores de λ . En el eje Y se muestran los valores antiguos y en el X, el resultado de la transformación. Tal y como se puede apreciar hay valores que contraen el crecimiento, por ejemplo $\lambda = -2$ y otros que lo aceleran, por ejemplo $\lambda = 3$.

El gráfico que aparece a continuación, recoge el efecto de las transformaciones para diferentes valores de λ . En el caso de que transformemos una variable X que tome valores positivos mayores que la unidad, con un valor de $\lambda < 1$, la variable transformada crecerá más lentamente que la original. Este efecto aumentará a medida que λ sea menor. En otras palabras, se están comprimiendo los valores de la variable transformada.

Por el contrario, si queremos que la variable que vamos a transformar tenga un crecimiento más rápido, será necesario utilizar valores de $\lambda > 1$. Mayores valores de λ implicarán un crecimiento más rápido.

Figura 5: Ejemplos de las transformaciones Box-Cox para valores de λ del -2 al 3 (Fuente: Scott, 1999)



3. Aplicaciones de la transformación Box-Cox

3.1 Una variable

A pesar de que con mucha frecuencia se asocian las transformaciones Box-Cox a problemas de Regresión Lineal, como se expondrá posteriormente, existen otras aplicaciones muy interesantes relacionadas con modificar una sola variable. En estos

casos, el objetivo que se desea cumplir puede ser, desde modificar la media o la varianza, a reducir la asimetría o la curtosis, pasando por el importantísimo problema de normalizar la variable.

Por ejemplo la importancia de la simetría se debe a:

- Permite saber cuál es el centro sin ambigüedades porque las medidas de tendencia central coinciden.
- Facilita la interpretación.
- Garantiza la robustez de algunos métodos estadísticos que requieren esta condición.

Para estudiar la asimetría se suelen utilizar fundamentalmente dos herramientas: el histograma y, como medida numérica el coeficiente de asimetría de Fisher (aunque existen otras medidas de asimetría).

Por otro lado, la importancia de la distribución normal se debe a (Batanero et al., 2001):

- Muchos fenómenos físicos, biológicos o psicológicos pueden ser modelizados mediante ella.
- Se considera también una adecuada aproximación a otras distribuciones como la T de Student, el modelo binomial o la Poisson (para determinados valores de sus parámetros).
- Esta distribución permite el cálculo de probabilidades de aparición de datos de esa distribución y por tanto a partir de una muestra de datos de una población se puede inferir datos de esa población.
- Para el uso de numerosos procedimientos estadísticos es necesario asumir la normalidad. Por ejemplo la ausencia desaconseja el uso de test paramétricos de contraste, o por ejemplo para el análisis de la varianza (ANOVA), regresión, procesos estocásticos, teoría de errores. Algunos autores como Zimmerman (1995) señalan que incluso contrastes no paramétricos pueden sufrir tanto como los paramétricos cuando el supuesto de normalidad no se cumple.

Para evaluar que los datos sigan una distribución normal, primero se recurre al QQ plot y después los contrastes de normalidad como por ejemplo el estadístico de Kolmogorov-Smirnov, Shapiro y Wilks, Jarque y Bera o (Avanza et al., 2002). En el caso de que la hipótesis de normalidad sea rechazada, sería necesario recurrir a las transformaciones.

Es importante destacar la utilización del método Box-Cox, porque además de ser una alternativa para detectar la normalidad y la linealidad, sirve también para transformar los datos que no lo sean (Peña y Peña, 1986), es decir, se obtiene una posible solución.

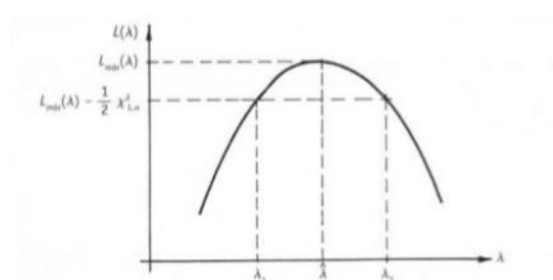
Algunos autores como Montgomery et al., (2002) consideran adecuado a la hora de estimar el valor de λ crear un intervalo de confianza tal y como queda reflejado en la Figura 6. Esto explica porque en algunos casos recogidos en este trabajo, a pesar de que el valor ajustado computacionalmente a través del Método de Máxima Verosimilitud que minimizaba la suma de cuadrados de los residuos era por ejemplo $\lambda=0,596$, el parámetro que finalmente se utilizaba para la transformación era $\lambda=0,5$. Una transformación de raíz es más sencilla de interpretar. Además, si en el intervalo de confianza se encuentra el valor de $\lambda=1$ no sería necesario realizar una transformación. El intervalo de confianza tendrá la siguiente forma (Hyde, 1999; Montgomery et al., 2001):

$$L(\lambda') - L(\lambda) \leq \frac{1}{2} X_{\alpha,1}^2 / n$$

Utilizar estos valores cercanos que corresponden a transformaciones típicas (véase Tabla 2) facilita los cálculos y la interpretación, pero también permite usar teoría ya conocida.

Figura 6: Ejemplo de estimación de λ por Máxima Verosimilitud incluyendo intervalo de confianza

(Fuente: Peña, 2001)



A pesar del valor que minimiza la suma de cuadrados de los residuos es λ' , los valores que se encuentran entre λ_1 y λ_2 podrían utilizarse de forma eficiente para realizar una transformación.

3.2 Dos variables

Al trabajar con dos variables, la transformación persigue objetivos distintos que pueden incluir también a los de los problemas de una variable. Entre ellos destacan ajustar la relación desconocida, conseguir la linealidad (es esencialmente equivalente a lo anterior) o trabajar con modelos de regresión lineal que cumplan todos los supuestos.

El modelo de regresión lineal se basa fundamentalmente en los supuestos de que el término error es aditivo, sigue una distribución normal, media igual a 0 y tiene una varianza constante (Rojo, 2007), además de ser independientes para datos distintos (cuando se supone normalidad, esto equivale a la incorrelación). Es necesario que se cumplan las siguientes propiedades:

$$E(\varepsilon) = 0$$

$$Var(\varepsilon) = \sigma^2$$

$$\varepsilon \sim N(0, \sigma^2)$$

En el caso concreto de datos temporales a estas características del error se le llama ruido blanco estricto: media igual a cero, varianza constante, incorrelado con su pasado y distribución normal (Vélez Serrano, 2006).

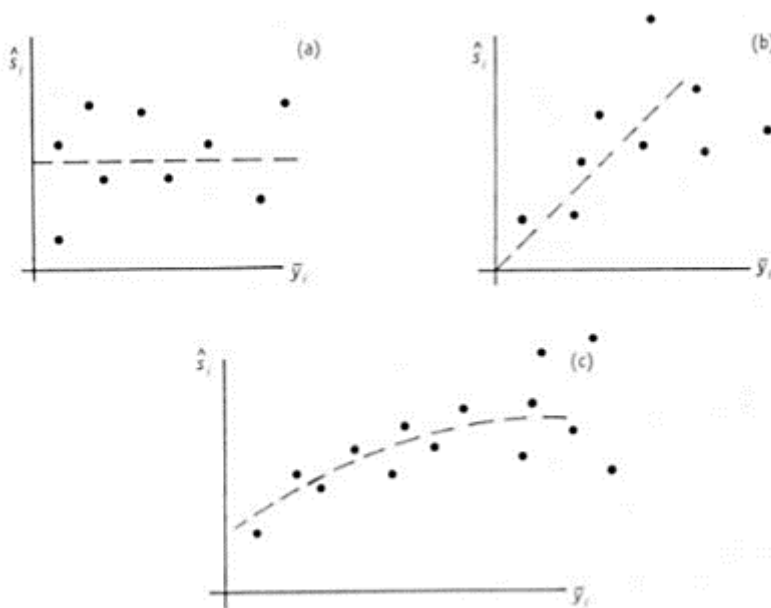
El problema de la varianza constante (para distintos valores de otra variable), hace referencia a la homocedasticidad. Si esta varianza no es homogénea, se dice que la serie tiene heterocedasticidad. Es muy importante corregir este error porque de otra manera, el término error no será impredecible, es decir, puramente aleatorio. Existen varios métodos para saber si existe heterocedasticidad, el más sencillo es observar la nube de puntos de los residuos. Aunque también se puede usar por ejemplo el contraste de White o el de Breusch-Pagan. De la misma manera existen contrastes de hipótesis para evaluar la autocorrelación, como por ejemplo el de Durbin-Watson o el de Box-Pierce.

A través de las transformaciones Box-Cox se puede intentar diseñar un modelo nuevo que preserve las características del original y que cumpla los supuestos:

- 1) Como la normalidad de la variable explicada Y es una hipótesis, en ocasiones se transforma esta variable de forma individual para conseguir esta normalidad.
- 2) Nada impediría transformar solo las variables explicativas, algunas o todas.
- 3) Por último, también es posible transformar todas las variables del modelo anterior.

Como consejo general: si los residuos del modelo ajustado tienen buenas propiedades no conviene transformar la variable explicada, mientras que en caso contrario, sí conviene transformar la variable explicada, y quizá, también las explicadas. Por otro lado, la relación entre los valores (en el gráfico) puede insinuar que valores de λ conviene probar. Afortunadamente, estos problemas se suelen resolver simultáneamente al transformar una o varias de las variables del modelo.

Figura 7: gráfico de residuos homocedásticos frente a heterocedásticos (Fuente: Peña, 2001).



En la Figura 7 se puede observar que en el gráfico (a) la distribución de los puntos que representan el término error es totalmente aleatoria. Por tanto existe homocedasticidad. Sin embargo, en los gráficos (b) y (c), se aprecia claramente como la distribución de la nube de puntos sigue un patrón y por tanto existe heterocedasticidad que conviene corregir. Por tanto para trabajar con datos que presentasen

heterocedasticidad sería conveniente realizar unas transformaciones que garantizase que la varianza fuese constante (Peña 2001).

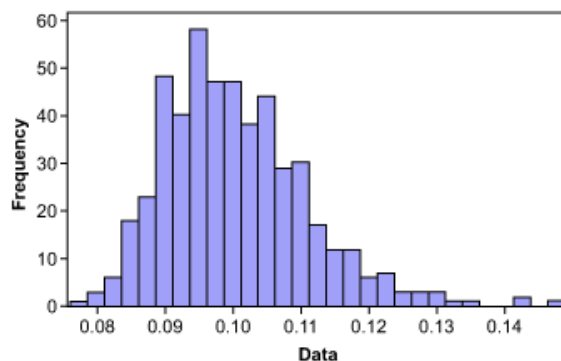
4. Ejemplos y ejercicios de simulación

Dado que el lector encontrará la transformación de Box-Cox en relación con problemas complejos como la heterocedasticidad o la regresión lineal (por ejemplo en libros de estadística y econometría), se incluyen aquí otro tipo de aplicaciones y ejemplos.

Ejemplo 1: Transformación Box-Cox datos asimétricos.

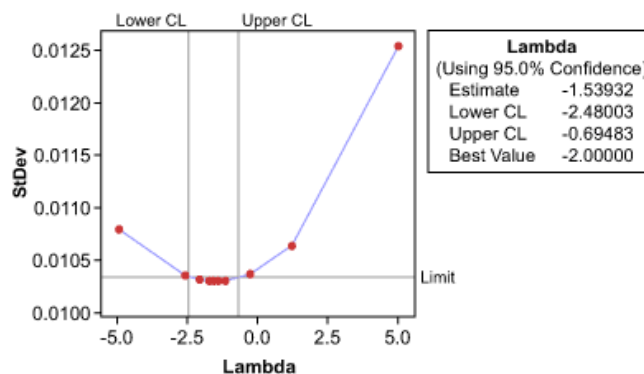
En este ejemplo concreto se pretende transformar datos que presentan asimetría. En el caso de que la ausencia de normalidad estuviese causada por este problema, al corregir la asimetría también se podría solucionar la normalidad.

Figura 8: Cycle Time, antes de la transformación (Fuente: Buthmann, visitado el 20/04/2017).



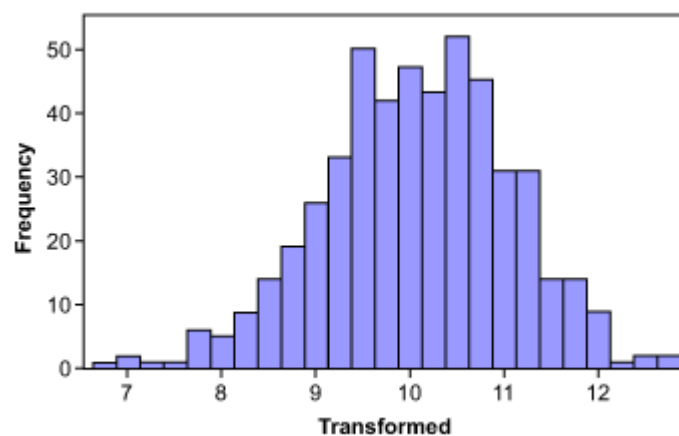
5.

Figura 9: Elección del parámetro λ para llevar a cabo la transformación (Fuente: Buthmann, visitado el 20/04/2017).



En la Figura 8, se observa que la distribución de los datos es asimétrica y por tanto para trabajar con ella sería adecuado realizar una transformación. Para la estimación del parámetro λ se ha construido un intervalo de confianza. En la Figura 9 se muestra cómo este intervalo de -2,48 a -0,69. El valor óptimo es -1,54 aunque la transformación es más sencilla si el número es redondeado y por tanto se hace uso de las transformaciones típicas (véase Tabla 2), facilitando así su estudio. En este caso concreto se trataría de -1 o de -2 y por tanto, dos posibles transformaciones del tipo $1/x^2$ o $1/x$ (véase Tabla 2).

Figura 10: Datos transformados usando $\lambda = -1$ (Fuente: Buthmann, visitado el 20/04/2017).

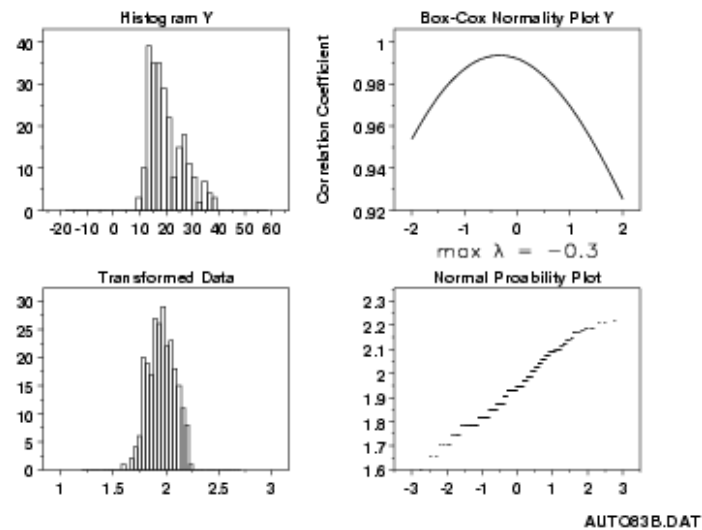


La Figura 10 muestra como al usar $\lambda = -1$, hay una mejora en la simetría y como consecuencia es probable que la distribución sea más próxima a la normal.

Ejemplo 2: Transformación Box-Cox con curtosis a la derecha.

La curtosis gráficamente no se puede observar. Para este ejemplo concreto el punto de partido son los cálculos realizados previamente por los autores que indican que efectivamente existe curtosis. Además gráficamente gracias al histograma se puede observar la ausencia de normalidad.

Figura 11: transformación Box-Cox con curtosis a la derecha (Fuente: NIST/SEMATECH, visitado el 1/05/2017).



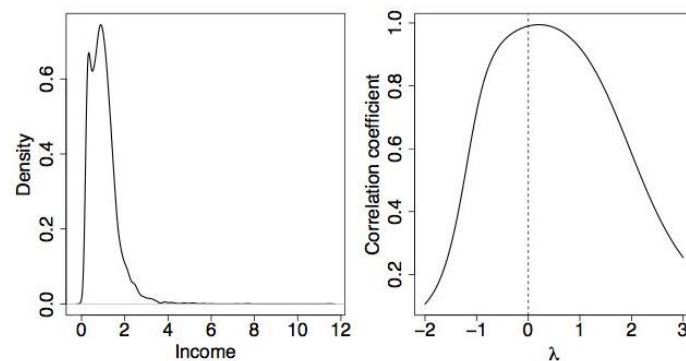
En la Figura 11, se aprecia como en el primer histograma de la variable Y presenta curtosis a la derecha y por tanto la muestra no sigue una distribución normal. El valor de λ elegido corresponde como se aprecia en el gráfico 'Box-Cox Normality plot Y' con el valor que garantiza el máximo punto de coeficiente de correlación, en este caso $\lambda = -0,3$.

En el segundo histograma, se observa la distribución de los datos después de aplicar la transformación Box-Cox, consiguiendo una distribución más normal que la previa.

Ejemplo 3: Transformación Box-Cox de datos de ingresos de las familias en 1973.

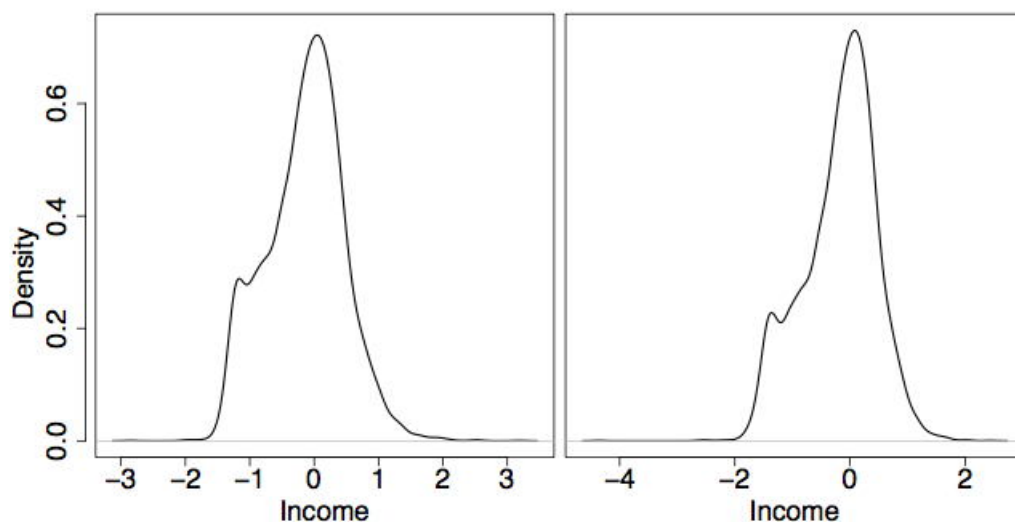
Este ejemplo muestra la transformación de datos de ingresos de las familias de Gran Bretaña durante el año 1973. Parte de una distribución como se aprecia en el histograma de la Figura 11 muy asimétrica.

Figura 12: Transformación de datos de ingresos de familias en Gran Bretaña, 1973. El primer gráfico muestra una gran asimetría. El segundo, el valor de λ que maximiza el coeficiente de correlación (Fuente: Scott, 1999).



La Figura 12 muestra dos gráficos, el primero donde se aprecia cómo los datos siguen una distribución asimétrica y el segundo, la estimación de λ a través de una maximización del coeficiente de correlación. Para este ejemplo el valor óptimo de λ es 0,21, lo que al redondear sugiere el valor $\lambda=0$, es decir, la transformación logarítmica.

Figura 13: Transformación de los datos para reducir la asimetría. En el primer gráfico se aplica $\lambda=0,21$ y en el segundo $\lambda=0$, es decir una transformación logarítmica (Fuente: Scott, 1999).



En la Figura 13 se muestra por un lado, el resultado de la transformación tras haber aplicado $\lambda=0,21$ y por otro lado, el resultado cuando $\lambda=0$. En ambos casos, a pesar de haber mejorado considerablemente, sigue existiendo algo de asimetría. El hecho de mostrar también la transformación para $\lambda=0$ es porque normalmente, a pesar de que $\lambda=0,21$ es el punto donde se maximiza la correlación, se suelen usar las transformaciones

típicas para facilitar el estudio y análisis (véase Tabla 2), como ya se ha indicado anteriormete.

Ejemplo 4: Transformaciones logarítmica frente a Box-Cox para niveles de polyamina en el plasma sanguíneo (Kutner et al., 2004).

La Tabla 4 recoge por un lado la edad x y los niveles de polyamina en el plasma sanguíneo de 25 niños. Como se puede apreciar en la tabla existe una mayor variabilidad entre los niños más pequeños con respecto a los mayores. Para un correcto análisis estadístico además de la construcción de un adecuado modelo de regresión lineal, es conveniente realizar una transformación. A menudo resulta difícil determinar cuál será la más apropiada para corregir a la vez la asimetría, la distribución de los residuos, las varianzas de los residuos y la no linealidad de la función de regresión.

Los valores del parámetro λ fueron de -1 a 1. El valor que minimizaba suma de errores al cuadrado (*sum of squared errors* SSE), fue el $\lambda=-0,5$ tal y como se aprecia en la Tabla 2. Para este ejemplo, los autores decidieron que al ser la transformación estable del 0 al -1, se usaría el logaritmo, es decir, $\lambda=0$ porque facilitaría la interpretación de los resultados obtenidos.

Tabla 3: Resultados de la transformación Box-Cox, elección del parámetro λ . (Fuente: Kutner et al., 2004)

λ	SSE	λ	SSE
1.0	78.0	-.1	33.1
.9	70.4	-.3	31.2
.7	57.8	-.4	30.7
.5	48.4	-.5	30.6
.3	41.4	-.6	30.7
.1	36.4	-.7	31.1
0	34.5	-.9	32.7
		-1.0	33.9

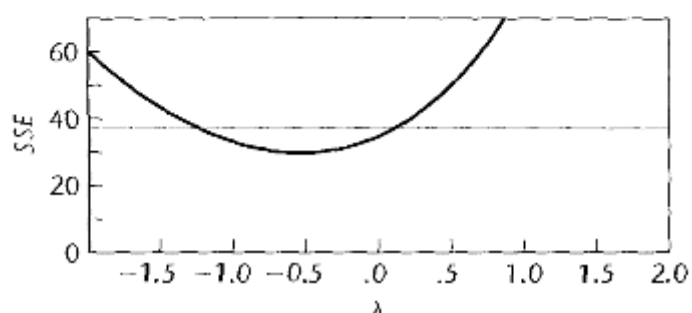


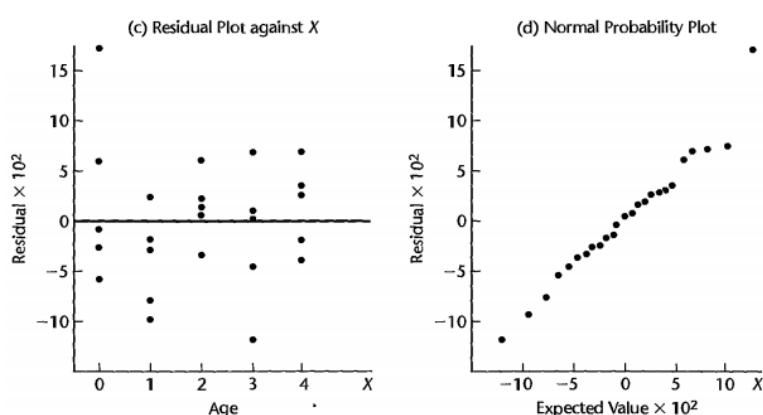
Tabla 4: Uso de las transformaciones logarítmicas para linealizar la regresión, estabilizar el error y la varianza en los niveles del plasma. (Fuente: Kutner et al., 2004)

Child i	(1) Age X_i	(2) Plasma Level Y_i	(3) $Y'_i = \log_{10} Y_i$
1	0 (newborn)	13.44	1.1284
2	0 (newborn)	12.84	1.1086
3	0 (newborn)	11.91	1.0759
4	0 (newborn)	20.09	1.3030
5	0 (newborn)	15.60	1.1931
6	1.0	10.11	1.0048
7	1.0	11.38	1.0561
...
19	3.0	6.90	.8388
20	3.0	6.77	.8306
21	4.0	4.86	.6866
22	4.0	5.10	.7076
23	4.0	5.67	.7536
24	4.0	5.75	.7597
25	4.0	6.23	.7945

La Tabla 4 muestra que tras aplicar la transformación logarítmica, la variabilidad de X se ha vuelto razonablemente constante además de haber conseguido establecer un modelo de regresión lineal tal que:

$$\hat{Y}' = 1,135 - 0,1023X$$

Figura 13: Gráfico de residuos y de normalidad (Fuente: Kutner et al., 2004)



La Figura 13 refleja cómo la transformación ha mejorado considerablemente la distribución de los errores y a través de un QQ plot se aprecia como los puntos se aproximan más a la línea bisectriz.

Ejemplo 5: datos con distribución asimétrica a los que se les aplican diferentes transformaciones para ver cuál se ajusta mejor a la distribución normal.

A continuación se incluye un ejercicio de simulación utilizando el programa estadístico Gretl. Se parte de unos datos que no siguen una distribución normal, tal y como se aprecia en la Figura 14, y a partir de ahí se han ido haciendo transformaciones seleccionando distintos valores de λ . La importancia de este ejemplo se refleja en que hoy en día gracias al uso de ciertos programas estadísticos, cualquiera puede probar las transformaciones.

Figura 14: Datos previos a la transformación.

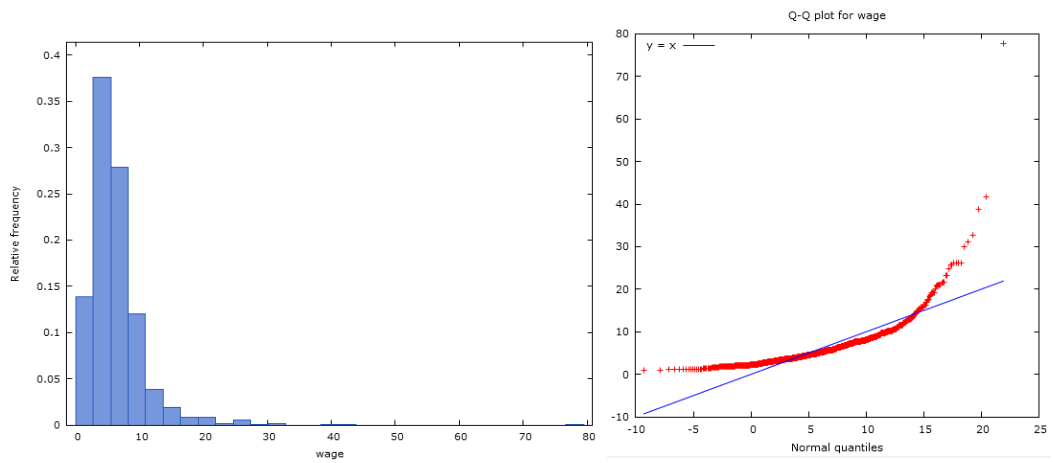


Figura 15: Transformación logarítmica: $\lambda=0$

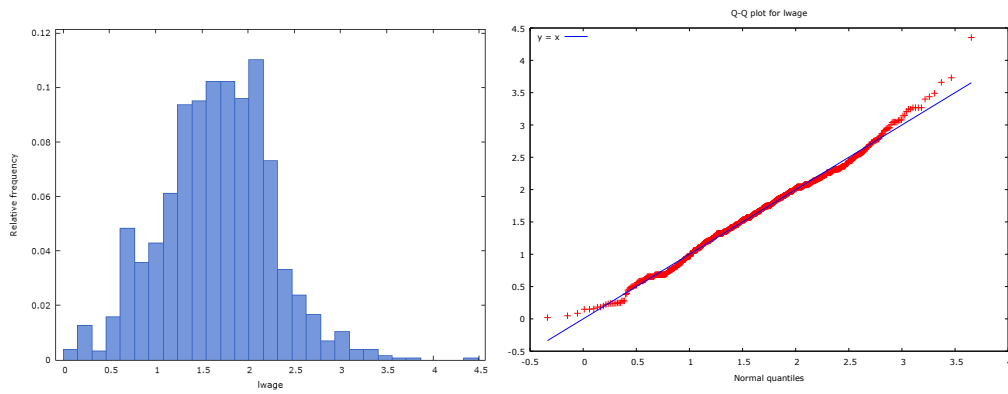


Figura 16: Transformación $\lambda=-0,25$

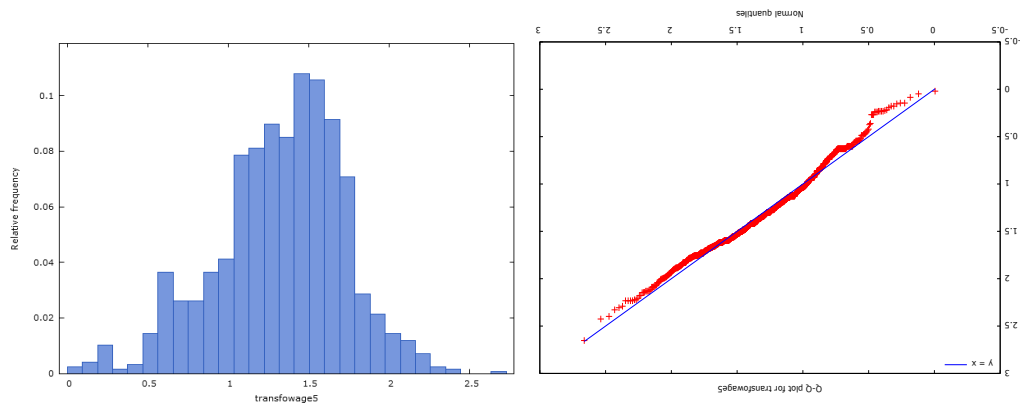


Figura 17: Transformación $\lambda=0,25$

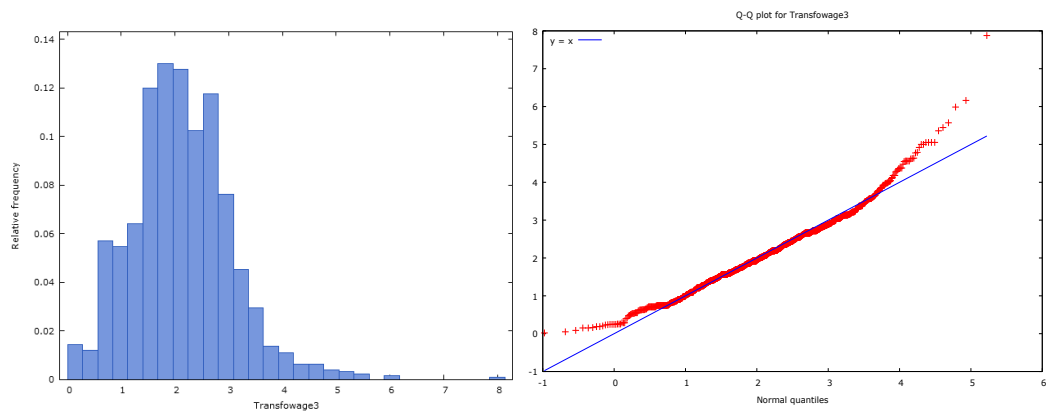


Figura 18: Transformación $1/\sqrt{x}$, $\lambda=0,5$

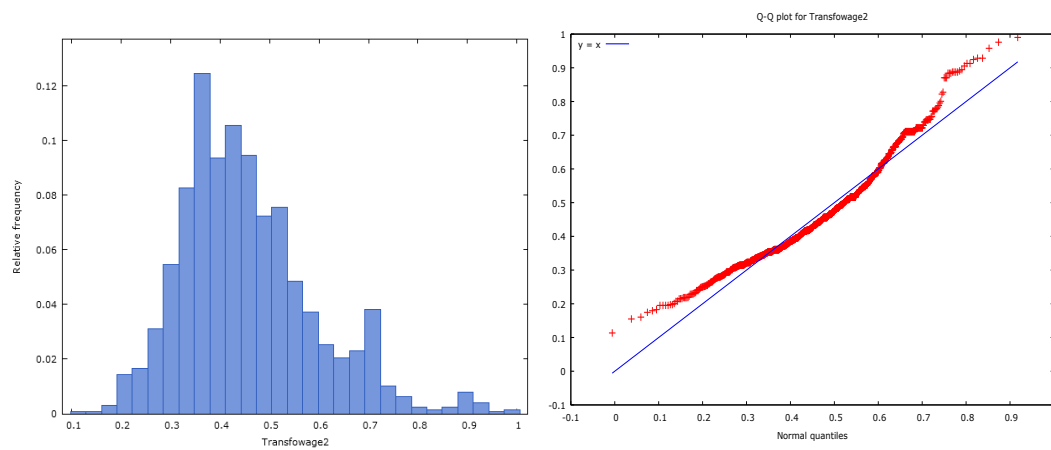


Figura 19: Transformación $\lambda=-0.1$

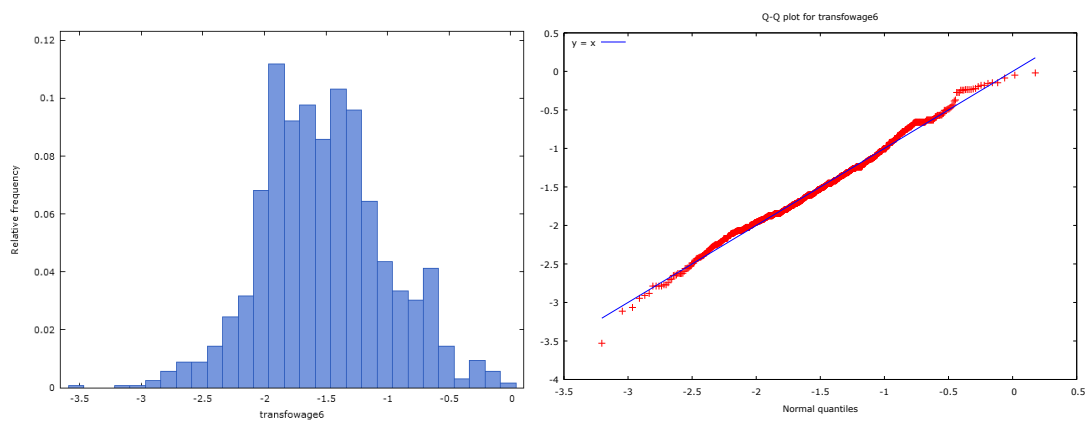
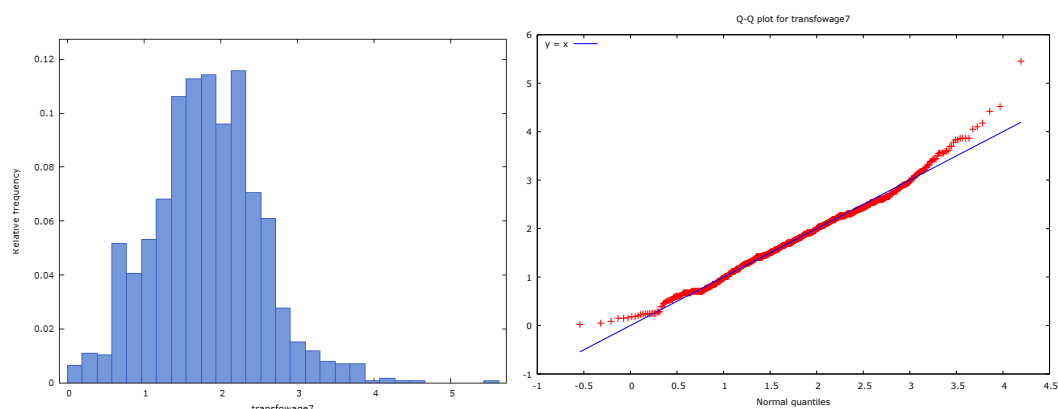


Figura 20: Transformación $\lambda=0,1$



En la Figura 14, se observa el histograma claramente asimétrico de unos datos antes de realizar cualquier tipo de transformación. El método que se ha utilizado es el de 'Prueba y error' y el programa estadístico Gretl.

Para la elaboración de este ejemplo el primer valor que se ha utilizado, ha sido tal y como se muestra en la Figura 15, $\lambda=0$, es decir, la transformación logarítmica. En esa misma figura a la derecha se observa en el QQ plot (herramienta estadística para contrastar la normalidad) como efectivamente la nube de puntos rojos se ajusta mucho mejor que antes a la bisectriz. Se puede apreciar como efectivamente, el uso de esta transformación ha mejorado considerablemente la distribución.

Sin embargo, la distribución todavía podría mejorar. Se han utilizado valores próximos al cero para ver si modificando poco a poco el valor de λ , la distribución se acercaba aún más a la normal. Los resultados han sido recogidos desde la Figura 16 a 20. El caso concreto de $\lambda=-0,1$ es el único con el que podría surgir la duda de si es más eficiente que $\lambda=0$. Aun así, tal y como se ha expuesto en el trabajo, el valor elegido sería $\lambda=0$ porque al ser una de las transformaciones típicas (véase Tabla 2) facilita el estudio y el análisis posterior.

La elaboración de este ejemplo ha reflejado además, como se cumplen las propiedades matemáticas de esta familia de transformaciones, especialmente la continuidad en función de λ (al ir variando λ poco a poco, la representación gráfica va 'cambiando lentamente').

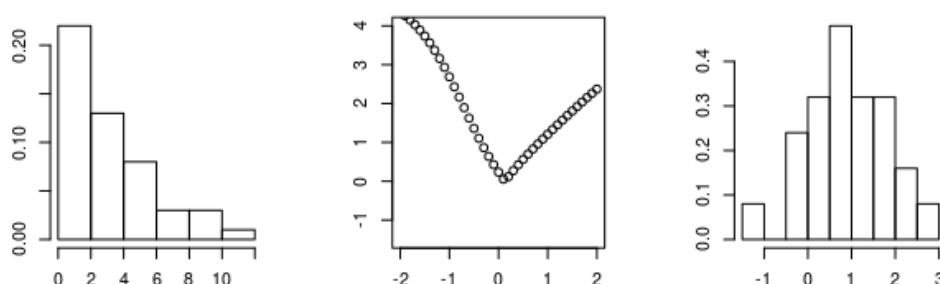
Ejemplo 6: ejercicio de simulación para una distribución X^2 con el objetivo de reducir la asimetría.

Gracias a las facilidades de cálculo que proporcionan los ordenadores, está al alcance de cualquier usuario programar un bucle que pruebe distintas transformaciones, evalúe el grado de cumplimiento del objetivo perseguido y elija la mejor. La evaluación mencionada depende del objetivo, y puede elegirse de forma más o menos formal. Por ejemplo, podría minimizarse una medida de asimetría, maximizarse el valor p de un contraste de normalidad (véase el ejemplo 7) o, más frecuentemente, maximizar la verosimilitud de la variable explicada de la regresión (lo que equivale a minimizar la varianza residual como indican Montgomery et al., 2002).

Para la realización de este ejemplo se ha utilizado el programa informático R. El primer paso ha sido crear un código (incluido en el Anexo 2) el cual permitía seleccionar el parámetro λ que minimice la distancia al 0. El lector podrá usarlo con otros datos para estimar cualquier otro valor de λ .

En la Figura 21 se aprecia como inicialmente el histograma de la distribución era asimétrico, como era de esperar, y tras aplicar una transformación con $\lambda=0,3$ mejora considerablemente la distribución de los datos.

Figura 21: Transformación Box-Cox $\lambda=0,3$, el antes y el después. (Fuente: elaboración propia).



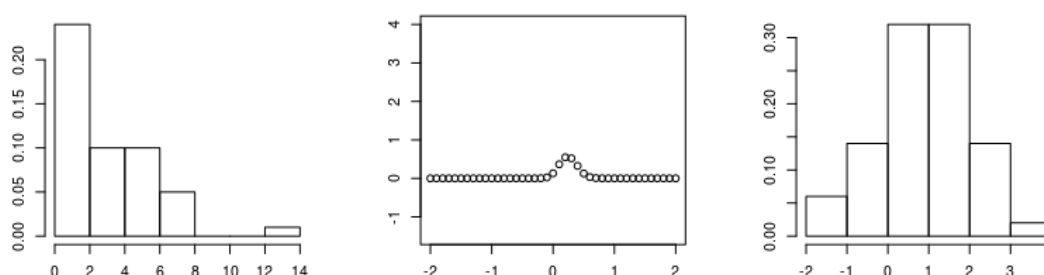
Se ha utilizado la distribución X^2 porque tiene una asimetría positiva como muchas variables económicas.

Ejemplo 7: ejercicio de simulación para una distribución X^2 con el objetivo de aumentar la normalidad.

Para la elaboración de este ejemplo de nuevo se ha empleado el programa informático R. El código está incluido en el Anexo 2 y el lector podrá utilizarlo con otros datos u otros valores del parámetro λ . Para este caso concreto el criterio para la elección del parámetro λ es maximizar el valor p al aplicar el test de Shapiro y Wilks.

En la Figura 22 se puede apreciar como inicialmente la distribución no era normal, y al aplicar una transformación en la que $\lambda=0,2$ la forma de la distribución mejora considerablemente. El valor óptimo, en esta ocasión como se dijo anteriormente y como se muestra en el gráfico es el valor que maximiza el valor p. En este caso el resultado del valor $p=0,89$. Por tanto, la hipótesis de normalidad es claramente aceptada.

Figura 22: Transformación Box-Cox $\lambda=0,2$, el antes y el después. (Fuente: elaboración propia).



5. Conclusiones

Las transformaciones tanto lineales como no lineales son útiles para hacer más fácil el análisis de los datos. Las primeras, aunque no permiten solucionar problemas como la ausencia de normalidad, la asimetría o la curtosis, pero si solucionan problemas como por ejemplo la comparación de datos en escalas diferentes. Las segundas, por el contrario, permiten reducir la asimetría o la curtosis, mejorar las características de normalidad de una variable o alcanzar el cumplimiento de las hipótesis de los modelos de regresión.

Los autores Box-Cox propusieron una familia de transformaciones que en principio permiten solucionar, parcial o totalmente, individual o simultáneamente los problemas

mencionados anteriormente. Para ello utilizan el parámetro λ , tal y como se ha explicado en este trabajo. Es importante destacar que su aplicación va más allá de las Ciencias Económicas y Empresariales; de hecho se ha incluido un ejemplo que hace referencia a la presencia de polyamina en el plasma sanguíneo.

Una vez que se ha estimado el parámetro λ a través de cualquiera de los métodos propuestos en la literatura (Máxima Verosimilitud, estimación Bayesiana y la técnica de probar distintos valores de λ), es fundamental comprobar que se han alcanzado los objetivos (los nuevos datos siguen efectivamente, una distribución normal, son homocedásticos etc.,) porque tal y como han apuntado numerosos autores posteriormente, no siempre es posible encontrar una transformación que garantice estos supuestos. También destaca el uso del propio contraste Box-Cox para detectar la normalidad.

Es necesario añadir, que a pesar de la estimación del parámetro λ , en la práctica la mayoría de los autores consultados para la elaboración de este trabajo tienden a emplear la transformación más típica que se aproxime a dicho valor. Para ello se crea un intervalo de confianza y así se estiman entre qué valores la transformación es eficiente (Hyde, 1999). Si en dicho intervalo se encuentran los valores de λ de los considerados típicos, normalmente se suelen usar estos últimos. Esto puede ser debido a que, para el estudio del comportamiento de los nuevos datos transformados es más sencillo emplear una transformación conocida, aunque no sea el valor estimado de λ más eficiente.

Por tanto emplear la familia de transformaciones Box-Cox tiene las siguientes ventajas: en primer lugar, se puede utilizar para contrastar la normalidad y a diferencia de otros contrastes indica cómo habría que transformar los datos (Peña y Peña, 1986). Es especialmente recomendable utilizarlo cuando el motivo de transformación es la asimetría, es decir, cuando se pretende detectar la ausencia de normalidad debida a la asimetría. En segundo lugar, las transformaciones típicas se suelen aplicar de forma intuitiva. Sin embargo, gracias a las transformaciones de Box-Cox se puede estimar el valor de λ que maximiza la eficiencia de la transformación, e incluso crear un intervalo de confianza para comprobar que valores de las transformaciones típicas están incluidos.

6. Anexos

Anexo 1: Demostración matemática transformaciones lineales

A continuación se incluye la demostración matemática que justifica que las transformaciones lineales cambian los valores numéricos de la variable, así como la media y la varianza, pero no la asimetría ni la curtosis.

Media:

$$E(ax + b) = a E(x) + b$$

Varianza:

$$Var(ax + b) = a^2 Var(x)$$

Asimetría y Curtosis

$$E \left[\left(\frac{(ax+b)-(a\mu+b)}{a\sigma} \right)^k \right] = E \left(\frac{a^k(x-\mu)^k}{a^k\sigma^k} \right) = E \left(\left(\frac{x-\mu}{\sigma} \right)^k \right)$$

Para la asimetría $k=3$ y para la curtosis $k=4$. En ambos casos se observa como el resultado final tras aplicarle una transformación del tipo lineal, corresponde con las propias fórmulas de las definiciones de asimetría y curtosis.

Anexo 2: Programación

Cómo hacer los cálculos: es necesario un bucle para ir utilizando los distintos valores del exponente (parámetro λ) incluidos en el vector e. Para cada valor:

- Es necesaria la familia de transformaciones que se quiere aplicar a los datos.
- Es necesaria otra función que evalúe el grado de cumplimiento del objetivo después de transformar los datos.
- Hace falta un bucle para ir seleccionando cada valor del parámetro. También hay que guardar el valor de la evaluación anterior.

- Finalmente, fuera del bucle es necesaria una función que aplique el criterio de optimización, minimizar o maximizar el grado de cumplimiento de los objetivos, e identifique la posición.
- Por último, se representan los histogramas antes y después de aplicar la transformación con el exponente seleccionado para comprobar gráficamente cómo ha cambiado la transformación.

Código

Muy importante: no utilizar los nombres 'x' y 'e' para ninguna variable nueva

DEFINICIONES

```
expo = function(lambda) x^lambda # Familia exponencial
expoT = function(lambda) x^lambda-1 # Familia exponencial trasladada
BoxCox = function(lambda) if (lambda!=0) (x^lambda-1)/lambda else log(x) # Familia de Box-Cox
geoMean = function(arg) exp(mean(log(arg))) # Media geométrica
geoBoxCox = function(lambda)# Familia de Box-Cox transformada
  if (lambda!=0) (x^lambda-1)/
    (lambda*(geoMean(x)^(lambda-1))) else geoMean(x)*log(x)
```

GRÁFICOS

```
x = seq(from=0.5, to=1.5, by=0.1) # Los valores o datos que se desea transformar
e = seq(from=-2, to=2, by=1) # Los exponentes o valores de lambda
```

Elegir una de estas cuatro familias

```
# Familia exponencial
f = expo

# Familia exponencial trasladada
f = expoT

# Familia de Box-Cox
f = BoxCox

# Familia de Box-Cox con la media geométrica
f = geoBoxCox
```



Elegir si representar una función al lado de otra o todas en un gráfico

Varias figuras (existe un máximo por línea)

```
x11()
par(mfcol=c(1, length(e)))
for (i1 in 1:length(e))
  plot(f(e[i1]), type='l', main="", xlab="", ylab="")
```

En una misma figura

```
x11()
k = c(-2,-1,0,1,2)
plot(f(e[1]), type='l', main="", xlab="", ylab="")
for (i1 in 1:length(e))
  lines(f(e[i1]))
```

Gráfico para comparar las cuatro familias de funciones

```
x11()
# Para seleccionar unos ejes verticales comunes a las cuatro figuras:
yLim = c(min(c(expo(e[1]), expoT(e[1]), BoxCox(e[1]), geoBoxCox(e[1]))),
          max(c(expo(e[1]), expoT(e[1]), BoxCox(e[1]), geoBoxCox(e[1])))))
for (i1 in 2:length(e)) {
  yLimTemp = c(min(c(expo(e[i1]), expoT(e[i1]), BoxCox(e[i1]), geoBoxCox(e[i1]))),
                 max(c(expo(e[i1]), expoT(e[i1]), BoxCox(e[i1]), geoBoxCox(e[i1])))))
  if (yLimTemp[1] < yLim[1])
    yLim[1] = yLimTemp[1]
  if (yLimTemp[2] > yLim[2])
    yLim[2] = yLimTemp[2]
}
```

Ahora hacemos las figuras

```
par(mfcol=c(1, 4))
plot(expo(e[1]), type='l', main="", xlab="", ylab="", ylim=yLim)
for (i1 in 1:length(e))
  lines(expo(e[i1]))
```

```

plot(expoT(e[1]), type='l', main="", xlab="", ylab="", ylim=yLim)
for (i1 in 1:length(e))
  lines(expoT(e[i1]))
plot(BoxCox(e[1]), type='l', main="", xlab="", ylab="", ylim=yLim)
for (i1 in 1:length(e))
  lines(BoxCox(e[i1]))
plot(geoBoxCox(e[1]), type='l', main="", xlab="", ylab="", ylim=yLim)
for (i1 in 1:length(e))
  lines(geoBoxCox(e[i1]))

```

EJERCICIOS

Elegir el problema de una variable que se quiere intentar resolver

Para reducir la asimetría

funTransfor = BoxCox # Transformación: Box-Cox

funEvalua = function(arg) mean(((arg-mean(arg))/sd(arg))^3) # Objetivo: asimetría

funOptimi = function(arg) which.min(abs(arg)) # Criterio: minimizar la distancia al 0

x = rchisq(50, 3) # Datos de una distribución con simetría positiva

Para aumentar la normalidad

funTransfor = BoxCox # Transformación: Box-Cox

funEvalua = function(arg) shapiro.test(arg)\$p.value # Objetivo: normalidad

funOptimi = function(arg) which.max(arg) # Criterio: maximizar el valor p

x = rchisq(50, 3) # Datos de una distribución no normal

Bucle que prueba valores de lambda y elige el que optimiza el objetivo

e = seq(-2, 2, 0.1) # Vector con los valores del parámetro lambda

vectorE = matrix(NA, length(e)) # Vector de evaluaciones

for (i1 in 1:length(e))

{

 datosTrans = funTransfor(e[i1])

 vectorE[i1] = funEvalua(datosTrans)

}

```

posLambdaEst = funOptimi(vectorE) # Posición del lambda
e[posLambdaEst] # Valor de lambda estimado

# Mostrar los resultados para la transformación elegida
x11()
par(mfcol=c(1, 3))
hist(x, freq=FALSE, labels=FALSE, main="", xlab="", ylab="")
plot(e, abs(vectorE), main="", xlab="", ylab="", ylim=yLim)
hist(funTransfor(e[lambdaEst]), freq=FALSE, labels=FALSE, main="", xlab="", ylab="")

```

7. Referencias

1. Avanza, M. Mazza, S. Martínez, G. Gimenez, L. (2002). Uso de transformaciones para el cumplimiento de los supuestos de normalidad y homocedasticidad, para modelos lineales.
2. Batanero, C. Tauber, L. M. Sánchez, V. (2001). Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos. *Cuadrante*, 10, 59-92.
3. Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26, 211-252.
4. Buthmann, A. (2017). Making Data Normal Using Box-Cox Power Transformation. *ISixSigma*. <https://www.isixsigma.com/tools-templates/normality/making-data-normal-using-box-cox-power-transformation/> (visitado 20/04/2017)
5. Carroll, R. J. 1980. A robust method for testing transformations to achieve approximate normality. *Royal Statistical Society, Series B*, 42, 71–78.
6. Fornieles, A. (2013). Transformaciones de datos en la elaboración de estudios salariales. *Revista Psicología del trabajo y de las Organizaciones*, 29 (2) 2174-0534.
7. Guisande González, C. Barreiro Felpeto, A. Maneiro Estraviz, I. Riveiro Alarcón, I. Vergara Castaño, A. R. Vaamonde Liste, A. (2006). Tratamiento de datos. *Editorial Díaz de Santos, S.A.*
8. Greene, W. H. (2002). *Econometric Analysis*. Prentice Hall
9. Hyde S. (1999). Likelihood based inference on the Box-Cox family of transformations: SAS and MATLAB programs.
10. John J. A., Draper N. R., (1980). An alternative Family of Transformations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 29 (2) 190-197
11. Kutner, M., Nachtsheim, C., Neter, J., and Li, W. (2004). *Applied Linear Statistical Models*, McGraw-Hill/Irwin, Homewood, IL.
12. Lane, D. M. (2010) Tukey's Honestly Significant Difference (HSD). In N. J. Elkind (Ed.) *Encyclopedia of Research Methods*, Sage Publications.
13. McCune & Grace (2002). *Ajustes de datos: transformación de datos*.

14. Montgomery, D., Peck E., y Vining C., (2002). Introducción al Análisis de Regresión Lineal. S.L. Grupo Paitria cultural, Alay Ediciones.
15. NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/> (visitado 20/04/2017)
16. Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation*, 8(6).
17. Peña, D. y Peña, J. I. (1986). Un contraste de normalidad basado en la transformación de Box-Cox. *Estadística española*, 110, 33-46.
18. Peña, D. (1997). Introducción a la estadística para las Ciencias Sociales. S.A. McGraw-Hill/Interamericana de España.
19. Peña, D. (2002). Regresión y diseño de experimentos. S.A. McGraw-Hill/Interamericana de España.
20. Ruppert, D. Aldershof B. Transformation to Symmetry and Homoscedasticity http://www.stat.ncsu.edu/information/library/mimeo.archive/ISMS_1760.pdf (visitado 20/04/2017)
21. Rojo Abuín, J. M. (2007) Regresión lineal múltiple. *Instituto de Economía y Geografía*.
22. Sakia, R. M. (1992) The Box-Cox transformation technique: a review. *The Statistician*, 41, 169-178.
23. Scott, D. Onlinestatbook. <http://onlinestatbook.com/lms/transformations/box-cox.html> (visitado 5/05/2017)



